We thank the referee for their detailed report. This is clearly a topic where there are differing views within the PDF community about certain issues, such as the appropriate use of a tolerance factor T and the relationship between a DIS-only and global PDF fit. It is quite clear that the referee has a rather different, strongly held, opinion with regards to some of these points in comparison to us. We are of course aware of this difference in approach, and have in fact tried hard to balance our discussion to avoid stating things too strongly with respect to either point of view, while trying to treat the (numerous) advantages of the LHeC and its relationship to LHC physics fairly. Moreover, we have paid special attention to understanding the differences that we find in comparison to previously reported studies, and indeed Sect 5 is fully devoted to such matters.

We are therefore disappointed that Referee 2 has not applied the same balance in approach when writing their report. While there are areas they highlight that certainly could be improved upon, in terms of the major comments these are all based on the unacceptable assumption that the referee's personal view on these matters is the only permitted one. We have included changes in some cases where the referee has helpfully pointed out areas where more discussion would be useful, but these major requests (i.e. asking us to add co–authors, to completely rewrite the text, sometimes entire sections, e.g. section 5, according to the referee's own personal view) are not ones that a referee should reasonably be making, as we hope that the SciPost editor-in-charge of our publication will understand.

The best we can do is therefore discuss here their points in detail, to refute the various rather strong statements they make, and justify again why the paper in something broadly following its current form should be published. We do this below (numbering according to the original report):

1. We have clearly referenced the source of the projections we use at the relevant place. We should of course point readers to the precise numbers, and this webpage is the only place that we are aware of that gives these. Moreover, we cite the corresponding papers as well as various talks on the subject throughout the paper. This is all done completely transparently and in a way that certainly does follow the usual scientific standard. We have added an additional citation to this in the introduction and to various figures/tables (as requested) to make the source of these projections clearer, but it not clear to us what other citations the referee would like us to provide in addition to the website (and indeed the referee does not specify).

   The comment about intellectual property is completely inappropriate and highly misleading. In all our communications with the scientists that produced the LHeC pseudo-data it has been communicated to us that the pseudo-data are publicly available for all interested parties, with the only requirement of acknowledging its origin, as we profusely do in the paper. To illustrate this, we have sent an example email to the editor–in–charge from Prof Max Klein (Liverpool), who produced the LHeC pseudo-data, inviting us to use it in our analyses (we do not add this to the publicly viewable webpage as it was a private communication). In all our communications with him it was clear that the data was openly available to the LHeC scientific community and that the only request is to acknowledge its origin, as we do. Moreover, exactly the same LHeC pseudo-data has been used in several other PDF projection studies (also cited in our manuscript) and the same policy has been adopted. One recent example is arXiv:1710.05935 (impact of LHeC on small-x BFKL studies): this paper provides important results that have been extensively discussed in the LHeC meetings, and no one there ever requested a shared authorship.

Therefore, the proposal that we invite the authors of the *ep* pseudo–data to be co–authors is inappropriate. As we discuss in the paper, the internal LHeC studies follow a completely different approach to ours, and clearly a co–authored paper would not be able to follow the approach we take. We consider it a useful contribution to this scientific exercise for us to be able to contribute our own independent study based on the same pseudo–data as the LHeC internal studies, and do not agree that co–authorship would be more credible or appropriate. The pseudodata are independent of any particular PDF projection study that one wishes to perform on them, and alternative studies based on these should be welcomed. We are sure that the referee agrees with us that the replication of scientific studies using independent methodologies cannot hamper scientific progress.

Finally, the concrete comment about 'mistakes' in our application of the pseudodata is in fact not justified; no such mistakes are present as far as we can tell. We will comment on this further below.

2. The comments here represent the referee's personal view on these matters, and it is simply incorrect to state these issues as matters of fact that everyone should follow, including ourselves in our paper. The descriptions of the future LHeC datasets as definitely 'self–consistent' cannot be assumed in such an unqualified way; as discussed below Fig. 4.1 there are known to be issues within the HERA–only data and corresponding $T = 1$ fits (tensions within the data, and with global fits in e.g. the up quark). For instance, it is know that it took several years of painful effort to combine the H1 and ZEUS cross-sections precisely because of the presence of tensions between the two sets of measurements, which had to be understood first. Clearly such issues of potential inconsistencies are generally larger in the case of hadron–hadron collisions, but they are not absent in the lepton–hadron case.

   Furthermore, the reason a more limited parameterisation is used in the current HERA–only cases is because these data place less tight constraints on the PDFs (flavour composition etc) and is not principally related to any self–consistency of the HERA data (though it is true that in global fits some of the parameterisation flexibility may be connected to inconsistencies in the data). The fact is that it would be very difficult (if not impossible) to pin down some aspects of PDFs, such the detailed quark flavour decomposition or the large-x gluon from a DIS–only fit.

   More importantly, it is a rather strong assumption to say that the very high precision LHeC data will be describable by the same rather restricted parameterisation, as is implied by the comment about using a different parameterisation for the LHeC (a point mentioned again in the referee's third comment). There is absolutely no reason to assume this will be the case. Moreover, as we discuss in detail in Section 5 the impact of this assumption can be very significant on the projected PDF uncertainties. It should really be stressed that this point is completely independent of whether one is considering a global fit or a DIS–only one; this is something we address in particular by reducing the prior (i.e. impact of the existing data in the fit) for the baseline sets.

   It is also important to be clear about what the referee's disagreement with the statement:

   In contrast to this understanding, the paper in its present from presents the current (PDF4LHC) and future (HL-LHC) global PDFs as one way to determine PDFs and only asks whether LHeC *would improve the precision.*

entails. We would precisely (and strongly!) argue that global PDF fits are one way to determine PDFs. Indeed, this idea underpins a great deal of LHC phenomenology. There is to us nothing controversial in the statement that global PDF fits provide a determination of PDFs, and that the LHeC can improve upon this. By arguing that one should only take data from the LHeC, and limit input from the HL–LHC to a consistency test, the referee is implicitly claiming that only DIS PDF fits can be trusted for collider phenomenology, and therefore that the range of precision LHC (and HL-LHC) measurements for PDF studies should be somehow discarded upon the arrival of the LHeC. This would also imply that our current theory predictions for example for Higgs production are simply unreliable or affected by much larger uncertainties than currently stated. While the referee is free to take this point of view, it should be emphasised that this viewpoint is not shared by the wide majority of the LHC physics community.

3. It is simply incorrect to state that we neglect future potential inconsistencies in *pp* data. This is precisely why a tolerance of $T = 3$ is taken. Now it is true that a greater degree of tension may appear in future data, requiring a larger tolerance, but then again it might not. We are quite open on this point and discuss it in detail on page 15. So we find it somewhat strange that the referee discusses this point about the tolerance as if it is not addressed in the paper. It is. Moreover, we show that the main qualitative findings of our paper are unchanged if different values of the tolerance are used.

   We are happy to add in a reference about the possibility of measuring the heavy quark masses, and the strong coupling, though this is not the focus of the current paper; this is addressed as discussed in the response to referee 3.

   Finally, as above, the statement that a tolerance of 1 is 'demanded' for the LHeC is an assumption, and moreover one that in the case of DIS–only fits from HERA is known to be a potentially questionable one. To restate, HERA–only fits are known to disagree with global fits within their $T = 1$ errors, and the HERA dataset itself is known to exhibit various tensions which tell us a matter of fact that the data/theory comparison is not following textbook statistics. Moreover, the fact that HERAPDF2.0 supplements the experimental PDF errors with model and parametrisation variations (which do not play a minor role) is an implicit admission that some kind of tolerance (understood as an increase of the PDF errors as compared to textbook statistics) is also required in their case. So we would argue that, on the contrary, there is currently no evidence that $T = 1$ is appropriate even for DIS-only fits.

Abstract

Discussing the missing jet data is clearly too detailed for the abstract, as is the point about theoretical precision. We are quite clear here and elsewhere that this is all done within a global PDF fit framework, where again a larger tolerance than 1 *must* be taken. So there is no inconsistency here.

Section 1

We have added a reference to arXiv:1802.04317, but otherwise unless the referee has something specific in mind, feel we have already provided quite a few references here. As we say above, we do comment on this point about interpreting a future e.g. high $p_\perp$ deviations, but agree the point about fitting away new physics could be made more clearly. As we discuss more below, we have added a more extended discussion about this.

We have added original references for Hessian profiling.

Section 2

As discussed above, this resource was provided to us upon request by the author of the pseudo-data, and we reference this all quite openly. There is simply no reason to say that this can only be used if the colleagues responsible for the pseudo-data are included as co–authors. We are quite clear and open about the method we take, and it is perfectly reasonable and standard scientific practice for us to provide an independent analysis based on the pseudo–data generated by these colleagues, without being required to follow their particular methodology for the PDF projections, or to include them as co–authors.

As far as we can see, all of the mistakes that the referee describes appear to be correct as they stand. The rapidity cut is clearly described in the pseudodata README (and indeed one can see from the pseudodata itself that a cut closely corresponding to this is imposed), as is the positron luminosity of 0.1 (and not 0.3) ab^{-1} (see also e.g. slide 51 of Claire Gwenlan's talk at DIS18). Possibly the referee is confusing this with the opposite polarization pseudodata on the latter point. $F_2^{c,\mathrm{CC}}$ is precisely the strange contribution that from what we understand the referee mistakenly believes we have not included. The last sentence on page 6 does not refer to the choice of parameterisation, but rather just to the pseudodata that is used, so is appropriate as it stands.

It is not clear what the referee would like in terms of understanding the flavour decomposition of PDF4LHC. This set is produced using a well understood procedure to combine various global fits. The parametric forms for these are described in detail in the respective papers, which we reference. To go into a discussion here of that would be unnecessary.

The comment about PDF4LHC ansatz being unsuitable for the LHeC pseudodata is not right, and the request for us to restrict ourselves to 'DGLAP–like' inputs misguided. As we have said above, the reason in e.g. the HERAPDF case that one takes a rather restricted form is because the data are not sensitive to e.g. the detailed quark flavour decomposition, and therefore one has to make an assumption here, fixing various parameters. The implication that all of the additional freedom that global fits include in the parameterisation in addition to this is due entirely (or even mainly for that matter) to the accommodation of inconsistencies in datasets is without basis. By far the dominant reason this extra freedom has been introduced as years go by is to account for the increasingly precise range of data, that pin down the individual PDF sets to increasing precision, and which could not be adequately captured by e.g. the HERAPDF parameterisation. There is only one 'true' PDF set in Nature: this fixing of parameters in the DIS–only case is a purely technical procedure to avoid an otherwise unstable fit with various redundant parameter directions if one allowed the parameterisation to be too flexible for the data constraints. It does not tell us anything about what this true form actually is, and moreover if we assume that a future very precise dataset as in the case of the LHeC will still respect these assumptions (even with some relatively mild loosening as in the most recent studies) it is liable to overestimate the data impact. This is all discussed in detail in Section 5.

The point about BFKL is mentioned elsewhere and is not relevant in the context of the current closure test.

Section 3

The comment about $T = 3$ being irrelevant for precision *ep* is, as we have discussed in detail above, questionable. Again, we are quite clear about our assumptions and indeed show results for both $T = 1$ and $T = 3$ choices, while discussing in detail the potential merits/pitfalls of both choices. There is no reason to insist that only $T = 1$ is used.

Section 4

We have included a figure showing the gluon PDF uncertainty on a linear scale.

Of course it is true that all HERA–only fits use $T = 1$, but the question is firstly how valid that is now, and more importantly how valid that will be in the future for LHeC–only fits. As we discuss in detail here there are numerous reasons that one could argue for a higher tolerance, even in a HERA–only fit. However we would also again point out that we are not firmly ruling out the lower tolerance, and indeed we present numerous results with $T = 1$ for precisely this reason, again providing a detailed discussion of all the relevant physics issues.

The reason we spell out these tensions in the HERA case is precisely because it is often stated in the literature that $T = 1$ is definitely the correct choice for DIS–only fits, as the referee does. We are also quite clear that there are various tensions in e.g. LHC data and that this is why a tolerance is used in global fits. This is all relevant for elucidating the discussion, and there is no good reason to remove this.

The point about Fig. 4.2 is misplaced. This is a closure test, so we assume the same theory and pseudodata- we could easily have included e.g. BFKL dynamics, but it is not relevant for the current study, which only focuses on PDF constraints. So there is no reason to highlight this point. There is also no reason to specify the tolerance we take in every figure caption: it is always $T = 3$, unless otherwise stated, as we discuss in the text.

The errors are normalized to the central value after profiling, so we do not believe the effect the referee describes will play a role here. We make clear at the bottom of section 2.2 that theory uncertainties are not included, though at NNLO precision for the LHC it is unclear that these would really have a large effect. In any case, there is no reason to spell that out in every figure caption.

We fit all currently publicly available pseudodata for the LHeC, and it is just as true that there are various HL–LHC data that we could consider in addition. We are completely clear on this point in the paper, and there is no need to restate it here. The point about new physics is discussed elsewhere and does not need to be restated again. We use a tolerance of 3 throughout, which is consistent with the global fit approach, and again the use of $T = 1$ is not without issues.

The point about D meson data is a fair one, and we have addressed this in line with the comment from referee 1.

Section 5

These comments from the referee are perhaps the most disappointing for us of all. The statements about our neglecting the recent developments are simply wrong: we clearly spell this out in the fourth paragraph of the section. The idea is then to study in a (yes) general way what impact our assumption about the PDF parameterisation can have when generating pseudodata with that parameterisation, i.e. the assumption that the future precise data will be describable within a certain parameterisation. We clearly show significant differences between the HERAPDF and more flexible PDF4LHC form. This is really completely unrelated to all of the issues above about global v.s. DIS–only fits, and the tolerance, which the referee has commented on at length, as we make explicit in Fig. 5.5. Within the context of a $T = 1$ LHeC–only projection, one will still get results that can be highly sensitive to the assumed parameterisation, and indeed such an effect has certainly been seen in various LHeC projections, with e.g. very small PDF uncertainties at very high and/or low x being clearly driven by parameterisation.

Now, we understand that efforts have been made to make the parameterisation more flexible within the LHeC internal studies, and indeed it is precisely for this reason that we have not directly compared to or made definitive comments in the paper about these projections. Nonetheless the overall parametric freedom used is still rather close to HERAPDF. It may well be true that the LHeC allows for a greater parametric freedom than HERA, and indeed this is precisely the point in that case: the parameterisation one uses in the projections needs to reflect this, and if not the corresponding PDF uncertainty reduction will be overstated. This is really an important point in all of this work, perhaps the most important one of the paper, and it cannot just be brushed aside.

Section 6

The point about new physics effects is definitely an important one; we have added a paragraph discussing this to the summary. The bottom line is that there are methods for disentangling such effects in a quantitative way at the LHC currently under active study. Nonetheless, this is a delicate issue and it is clearly a particular advantage of the LHeC that such effects should be largely absent. We have emphasised this in the new addition to the text.